

# A transference principle for Ramsey numbers of bounded degree graphs

Choongbum Lee <sup>\*</sup>

## Abstract

We investigate Ramsey numbers of bounded degree graphs and provide an interpolation between known results on the Ramsey numbers of general bounded degree graphs and bounded degree graphs of small bandwidth. Our main theorem implies that there exists a constant  $c$  such that for every  $\Delta$ , there exists  $\beta$  such that if  $G$  is a graph with maximum degree at most  $\Delta$  having a homomorphism  $f$  into a graph  $H$  of maximum degree at most  $d$  where  $|f^{-1}(v)| \leq \beta n$  for all  $v \in V(H)$ , then the Ramsey number of  $G$  is at most  $c^{d \log d} n$ . A construction of Graham, Rödl, and Ruciński shows that the statement above holds only if  $\beta \leq (c')^\Delta$  for some constant  $c' < 1$ . We further study the parameter  $\beta$  using a density-type embedding theorem for bipartite graphs of small bandwidth. This theorem may be of independent interest.

## 1 Introduction

The *Ramsey number* of a graph  $G$ , denoted  $r(G)$ , is the minimum number  $n$  such that every edge-coloring of  $K_n$  using two colors admits a monochromatic copy of  $G$ . It was first studied in the seminal paper of Ramsey [24] which established that the Ramsey number of the complete graph  $K_k$  on  $k$  vertices is finite for all positive integers  $k$ . Since then, Ramsey theory, the study of various results that can be grouped under the common theme “every large system has a well-organized subsystem”, flourished and became one of the most active fields of research in combinatorics. It is a beautiful field with many questions still remaining to be answered and has deep connections to other fields such as logic, geometry, and computer science. See the classical book of Graham, Rothschild, and Spencer [17] for a comprehensive overview of the field, or a survey of Conlon, Fox, and Sudakov [10] for recent developments in graph Ramsey theory.

In this paper we study the Ramsey number of bounded degree graphs. The history of such study can be traced back to a paper of Burr and Erdős [6] from 1975 which predicted that the behavior of Ramsey numbers of sparse graphs will be dramatically different from that of the complete graph (the Ramsey number of complete graphs is exponential in terms of the number of vertices [13, 12]). A graph  $G$  is  $d$ -*degenerate* if all its subgraphs has a vertex of degree at most  $d$ . In their paper, Burr and Erdős conjectured that for all  $d$ , there exists a constant  $c = c(d)$

---

<sup>\*</sup>Department of Mathematics, MIT, Cambridge, MA 02139-4307. Email: cb\_lee@math.mit.edu. Research supported by NSF Grant DMS-1362326.

such that  $r(G) \leq c(d)n$  for all  $n$ -vertex  $d$ -degenerate graphs  $G$ . This conjecture is still open (see [19, 22]).

In 1983, Chvátal, Rödl, Szemerédi, and Trotter [7] showed that the Burr-Erdős conjecture holds if the degeneracy condition is replaced with a bounded degree condition. More precisely, they showed that for all  $\Delta$ , there exists a constant  $c = c(\Delta)$  such that  $r(G) \leq c(\Delta)n$  for all  $n$ -vertex graphs  $G$  of maximum degree at most  $\Delta$ . Their proof relied on the regularity lemma and gave a tower-type dependency between  $c(\Delta)$  and  $\Delta$ . Since then, the bound on  $c(\Delta)$  has been improved by Eaton [11], Graham, Rödl, and Ruciński [15, 16], and by Conlon, Fox, and Sudakov [9] who showed that there exists a constant  $c$  such that  $r(G) \leq c^{\Delta \log \Delta} n$  holds for  $n$ -vertex graphs  $G$  of maximum degree at most  $\Delta$ . For bipartite graphs, independently, Conlon [8], and Fox and Sudakov [18] showed that a better bound  $r(G) \leq c^\Delta n$  holds (for some other constant  $c$ ). On the other hand, Graham, Rödl, and Ruciński [16] showed that there exists a constant  $c > 1$  such that for all large enough  $n$ , there exists a bipartite graph  $G$  with maximum degree at most  $\Delta$  satisfying  $r(G) \geq c^\Delta n$ .

In some cases the constant is known to be significantly smaller. The *bandwidth* of an  $n$ -vertex graph  $G$  is the minimum integer  $b$  for which there exists a labelling of the vertices by  $[n]$  such that  $|i - j| \leq b$  holds for all edges  $\{i, j\} \in E(G)$ . Allen, Brightwell, and Skokan [1] showed that if  $G$  is an  $n$ -vertex with maximum degree at most  $\Delta$  and bandwidth at most  $\beta n$ , then  $r(G) \leq (2\chi(G) + 4)n \leq (2\Delta + 6)n$ . Despite the rather wide gap between the two constants  $c^\Delta$  and  $2\Delta + 6$ , not much is known about the constant that dictates the Ramsey number of graphs of bounded maximum degree. Since it is known that a graph has small bandwidth if and only if it has poor expansion property (see [3] for more detail) and the example of Graham, Rödl, and Ruciński is a good expander, one can reasonably guess that the constant  $c_G$  for which  $r(G) \leq c_G \cdot n$  depends on the expansion property of  $G$ . In this paper, we study the relation between the constant  $c_G$  and the structure of the graph  $G$  in further depth.

Throughout the paper, when considering 2-edge-coloring of a graph, we tacitly assume that the two colors are red/blue, respectively, and refer to the subgraph consisting of the red edges as the *red graph*, and of the blue edges as the *blue graph*. A (*vertex*) *weighted graph* is a pair  $(G, w)$  of a graph  $G$  and a weight function  $w : V(G) \rightarrow [0, 1]$ . For a set  $X \subseteq V(G)$ , define  $w(X) = \sum_{x \in X} w(x)$ . For two graphs  $G$  and  $H$ , a *homomorphism* from  $G$  to  $H$  is a map  $f : V(G) \rightarrow V(H)$  such that  $\{f(v), f(w)\} \in E(H)$  whenever  $\{v, w\} \in E(G)$ .

**Definition 1.1.** *The Ramsey number of a weighted graph  $(G, w)$ , denoted  $\hat{r}(G, w)$  is the minimum integer  $n$  satisfying the following: for every 2-edge-coloring of  $K_n$ , there exists a homomorphism  $f$  from  $G$  to the red graph, or to the blue graph, for which  $w(f^{-1}(v)) \leq 1$  holds for all  $v \in V(K_n)$ . We simply denote  $\hat{r}(G, w)$  as  $\hat{r}(G)$  when the weight function is clear from the context.*

Note that the Ramsey number of weighted graphs generalizes the Ramsey number of graphs since given a graph we can always consider a constant weight function assigning weight one to all vertices. For this weight function, the Ramsey number in the traditional sense is equal to the Ramsey number as defined above. This observation in particular implies that  $\hat{r}(G, w)$  is finite

for all weighted graphs  $(G, w)$ . Further note that if  $G$  is the complete graph, then for all weight functions  $w$ , the Ramsey number of  $(G, w)$  equals the Ramsey number of  $G$  since a homomorphism from a complete graph to a graph with no loops is necessarily injective. On the other hand suppose that  $G$  is  $k$ -colorable and suppose that  $w$  is a weight function where  $w(X) \leq 1$  for each of the  $k$  color classes  $X$  of  $G$ . Then one can easily check that  $\hat{r}(G) \leq r(K_k) \leq \binom{2k}{k}$ . Therefore both the structure of  $G$  and the weight function  $w$  plays an important role in determining the Ramsey number of weighted graphs. However we will later see that for bounded degree graphs  $G$ , the Ramsey number of  $(G, w)$  is mostly determined by the total weight  $w(V(G))$  of the graph.

We consider another generalization of Ramsey numbers, implicitly studied in [1], where the host graph is a graph of large minimum degree instead of the complete graph.

**Definition 1.2.** *For a positive real  $\varepsilon$  and a weighted graph  $(G, w)$ , define the  $\varepsilon$ -stable Ramsey number  $\hat{r}_\varepsilon(G, w)$  as the minimum integer  $n$  satisfying the following: for every graph  $\Gamma$  on  $n$  vertices of minimum degree at least  $(1 - \varepsilon)n$ , for every 2-edge-coloring of  $\Gamma$ , there exists a homomorphism  $f$  from  $G$  to the red graph, or to the blue graph, for which  $w(f^{-1}(v)) \leq 1$  holds for all  $v \in V(\Gamma)$ .*

The stable Ramsey number generalizes Ramsey number since  $\hat{r}(G, w) = \hat{r}_\varepsilon(G, w)$  holds for every weighted graph  $(G, w)$  if  $\varepsilon < \frac{1}{\hat{r}(G, w) - 1}$ . However, given a weighted graph  $(G, w)$ , the  $\varepsilon$ -stable Ramsey number does not necessarily exist. For example if  $G$  is an  $r$ -partite graph, then  $\hat{r}_\varepsilon(G, w)$  does not exist for  $\varepsilon \geq \frac{1}{r-1}$  since we can take the host graph  $\Gamma$  to be a complete  $(r-1)$ -partite graph. In fact  $\hat{r}_\varepsilon(G, w)$  is finite if and only if  $\varepsilon < \frac{1}{r(K_{X(G)}) - 1}$  (see Section 6). The following theorem extends a theorem of Conlon, Fox, and Sudakov and shows that for bounded degree graphs, the stable Ramsey number is mostly determined by the total weight of the graph.

**Theorem 1.3.** *There exist constants  $c$  such that the following holds for every natural number  $\Delta$  and positive real number  $\varepsilon$  satisfying  $\varepsilon < c^{-\Delta \log \Delta}$ . If  $(G, w)$  is a weighted graph with maximum degree at most  $\Delta$ , then  $\hat{r}_\varepsilon(G) \leq c^{\Delta \log \Delta} \cdot w(V(G))$ .*

The following result is the main theorem of this paper studying the Ramsey number of bounded degree graphs. It roughly asserts that if  $G$  is a subgraph of a blow-up of  $H$ , then the Ramsey number of  $G$  can be described in terms of the Ramsey number of  $H$ .

**Theorem 1.4.** *For all  $\Delta, \xi$  and  $\varepsilon$ , there exists  $\beta$  and  $n_0$  such that the following holds for all  $n \geq n_0$ . Let  $G$  and  $H$  be graphs where  $G$  has  $n$  vertices and maximum degree at most  $\Delta$ . Suppose that there exists a homomorphism  $f$  from  $G$  to  $H$  for which  $|f^{-1}(v)| \leq \beta n$  for all  $v \in V(H)$ . Then for the weight-function of  $H$  defined by  $w(v) = \frac{1}{\beta n} |f^{-1}(v)|$  we have  $r(G) \leq (1 + \xi) \hat{r}_\varepsilon(H, w) \cdot \beta n$ .*

Note that the weight function  $w$  defined in Theorem 1.4 satisfies  $w(V(H)) = \frac{1}{\beta}$ . A *wheel graph*  $W_k$  is a graph with  $k$  vertices consisting of a cycle on  $k-1$  vertices and a vertex adjacent to all vertices on the cycle. In Section 3, we will see that there exist constants  $\varepsilon$  and  $c$  such that  $\hat{r}_\varepsilon(W_k, w) \leq c \cdot w(V(W_k))$  for all  $k$  and all weight functions  $w : V(W_k) \rightarrow [0, 1]$ . Hence if  $G$  has maximum degree at most  $\Delta$  and a homomorphism  $f$  into a wheel graph  $W_k$  where  $|f^{-1}(v)| \leq \beta n$  for all  $v \in V(W_k)$ , then by Theorem 1.4 above with  $\xi = 1$ , we obtain  $r(G) \leq 2\hat{r}_\varepsilon(W_k, w) \cdot \beta n \leq 2cn$

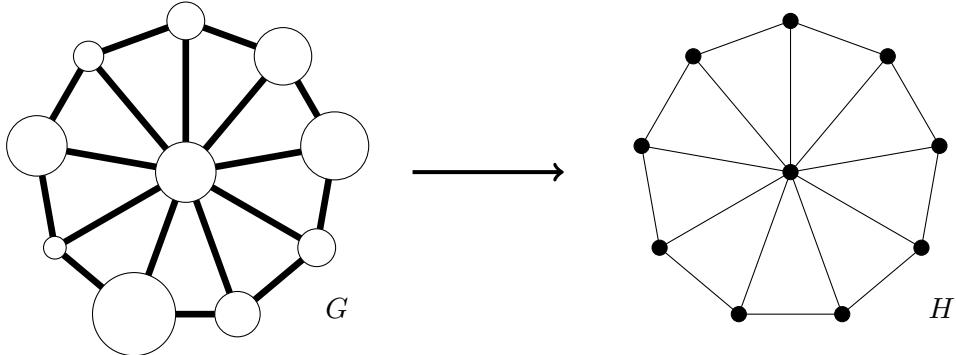


Figure 1: A graph of maximum degree at most  $\Delta$  and a homomorphism into a wheel graph.

(see Figure 1). Note in particular that the constant does not depend on  $\Delta$ . This is in sharp contrast with the bound  $r(G) \leq c^{\Delta \log \Delta} n$  (the constant  $c$  is different from above) that we obtain through the theorem of Conlon, Fox, and Sudakov. As another example, if  $H$  has maximum degree at most  $d$ , then by Theorem 1.3 we see that  $\hat{r}_\varepsilon(H, w) \leq c^{d \log d} \frac{1}{\beta}$  for small enough  $\varepsilon$ . By applying Theorem 1.4 with the  $\xi = 1$  and  $\varepsilon < c^{d \log d}$ , we obtain the following corollary.

**Corollary 1.5.** *There exists a constant  $c$  such that for all  $\Delta$ , there exists  $\beta$  and  $n_0$  such that the following holds for all  $n \geq n_0$ . Let  $G$  be a  $n$ -vertex graph with maximum degree at most  $\Delta$ , and  $H$  be a graph with maximum degree at most  $d$ . Suppose that there exists a homomorphism  $f$  from  $G$  to  $H$  for which  $|f^{-1}(v)| \leq \beta n$  for all  $v \in V(H)$ . Then  $r(G) \leq c^{d \log d} n$ .*

It is known (implicitly in [4]) that for all  $r$ , there exists a constant  $c > 1$  such that if  $G$  is an  $r$ -partite graph of bandwidth at most  $\beta n$ , then there exists a homomorphism  $f$  from  $G$  to the  $r$ -th power of a path of length  $\frac{1}{c\beta}$  where  $|f^{-1}(v)| \leq c\beta n$  for all  $v$ . Allen, Brightwell, and Skokan's result  $r(G) \leq (2\chi(G) + 4)n$  mentioned above then follows from Theorem 1.4 and a bound on the Ramsey number of power of paths (in fact the proof of Theorem 1.4 is based on a generalization of their proof).

The necessity of forcing  $|f^{-1}(v)| \leq \beta n$  for some small constant  $\beta$  can be seen from the example of Graham, Rödl, and Ruciński. They proved that there exists a constant  $c < 1$  such that for all  $\Delta$  and large enough  $n$ , there exists a  $c^\Delta n$ -vertex bipartite graph  $G$  of maximum degree at most  $\Delta$  for which  $r(G) > n$ . If  $\beta \geq 20c^\Delta$  in Theorem 1.4, then there exists a homomorphism  $f$  from  $G$  to  $K_2$  such that  $|f^{-1}(v)| \leq \frac{1}{20}\beta n$  for both vertices  $v$  of  $K_2$ . Since  $\hat{r}_\varepsilon(K_2, w) = 2$ , Theorem 1.4 (if true) will imply that  $r(G) \leq (1 + \xi)2\beta n < n$  which is a contradiction. Thus we see that  $\beta$  must be at most  $20c^\Delta n$  in Theorem 1.4. On the other hand, the bound on  $\beta$  that we obtain in Theorem 1.4 has a tower-type dependency on  $\Delta$ . It would be interesting to determine the best possible value of  $\beta$  that we can take. The following density embedding theorem has an interesting implication towards this problem.

**Theorem 1.6.** *Let  $G$  be an  $n$ -vertex graph of minimum degree at least  $(\delta + \alpha)n$ . Then  $G$  contains all bipartite graphs  $H$  on at most  $\delta n$  vertices with maximum degree at most  $\Delta$  and bandwidth at*

most  $\frac{1}{256\Delta}\alpha^{6\Delta+1}n$ .

Theorem 1.6 can be seen as an extension of density embedding theorems of bipartite graphs proved by Conlon [8], and Fox and Sudakov [18], and may be of independent interest. Note that Theorem 1.6 is asymptotically tight in terms of the number of vertices of  $H$ . As we will later see (Corollary 5.2), it implies that  $r(G) \leq (4 + \varepsilon)n$  if  $G$  is a bipartite graph with maximum degree at most  $\Delta$  and bandwidth at most  $c^\Delta n$  for some positive constant  $c < 1$ . Similar result can be obtained by the theorem of Allen, Brightwell, and Skokan but with a worse bound on the bandwidth. This corollary implies that a transference-type result holds even when  $\beta$  is as large as  $c^\Delta n$  for the special case when  $G$  is a bipartite graph with small bandwidth.

A  $d$ -dimensional hypercube  $Q_d$  is a graph with vertex set  $\{0, 1\}^d$  where two vertices are adjacent if and only if they differ in exactly one coordinate. Since the bandwidth of  $Q_d$  is known to be  $O(\frac{2^d}{d})$ , Theorem 1.6 is closely related to another conjecture of Burr and Erdős stating that there exists a constant  $c$  such that  $r(Q_d) \leq c2^d$  holds for all natural numbers  $d$ . Unfortunately the bandwidth condition in Theorem 5.2 is too weak to imply the conjecture.

The rest of the paper is organized as follows. In Section 2 we prove Theorem 1.4 using a variant of the blow-up lemma whose proof we defer to a later section. In Section 3 we establish a bound on the weighted Ramsey number of wheel graph and prove Theorem 1.3. In Section 4 we prove the variant of the blow-up lemma used in Section 2. In Section 5, we prove Theorem 1.6 and then conclude with some remarks in Section 6.

**Notation.** A graph  $G = (V, E)$  is given by a pair of vertex set  $V$  and edge set  $E$ . For a vertex  $v \in V$ , define  $\deg(v)$  as the degree of  $v$ , and for a set  $X \subseteq V$ , define  $\text{codeg}(X)$  as the number of common neighbors of the vertices in  $X$ . For two vertices  $v, w \in V$ , we define  $\text{codeg}(v, w) = \text{codeg}(\{v, w\})$ . For a set  $X \subset V$ , define  $G[X]$  as the subgraph of  $G$  induced on  $X$ . For a pair of sets  $X, Y \subseteq V$ , define  $e(X, Y)$  as the number of pairs  $(x, y) \in X \times Y$  that form an edge in  $G$ . When  $X, Y$  are disjoint sets, define  $d(X, Y) = \frac{e(X, Y)}{|X||Y|}$ . When there are several graphs under consideration, we often use subscript such as in  $e_G(X, Y)$  to clarify the graph that we are referring to. For two graphs  $H$  and  $G$ , an *embedding* of  $H$  to  $G$  is an injective map  $f : V(H) \rightarrow V(G)$  for which  $\{f(v), f(w)\} \in E(G)$  whenever  $\{v, w\} \in E(H)$ . An *embedding* of a weighted graph  $(H, w)$  into a graph  $G$  is a map  $f : V(H) \rightarrow V(G)$  for which  $\{f(v), f(w)\} \in E(G)$  whenever  $\{v, w\} \in E(H)$  and  $w(f^{-1}(v)) \leq 1$  for all  $v \in V(H)$ . For a set  $X \subseteq V(H)$ , a *partial embedding on  $X$*  is an embedding of  $H[X]$  to  $G$ . For a finite set  $X$  and a natural number  $n$ , we use the notation  $X^n$  to denote the product space  $X \times X \times \dots \times X$  where product is taken  $n$  times. Equivalently,  $X^n$  is the set of ordered  $n$ -tuples of elements of  $X$ .

We use  $\log$  without subscript to denote base 2 logarithm. We omit floor and ceiling whenever they are not crucial. Throughout the paper, we use constants with subscripts such as in  $\beta_{2.3}$  to indicate that  $\beta$  is the constant coming from Theorem/Corollary/Lemma/Proposition 2.3.

## 2 Transference principle

Let  $G$  be a graph on  $n$  vertices. A pair of disjoint vertex subsets  $(X, Y)$  is  $\varepsilon$ -regular if for all  $X' \subseteq X$  and  $Y' \subseteq Y$  satisfying  $|X'| \geq \varepsilon|X|$  and  $|Y'| \geq \varepsilon|Y|$ , we have  $|d(X, Y) - d(X', Y')| \leq \varepsilon$ . A partition  $V(G) = V_0 \cup V_1 \cup \dots \cup V_k$  is  $\varepsilon$ -regular if (i)  $|V_0| \leq \varepsilon n$ , (ii)  $|V_i| = |V_j|$  for all  $i, j \geq 1$ , and (iii) for each  $i \in [k]$ , there exists at most  $\varepsilon k$  indices  $j \in [k]$  for which  $(V_i, V_j)$  is not  $\varepsilon$ -regular.<sup>1</sup> We define the  $\varepsilon$ -reduced graph of a partition  $\{V_i\}_{i=0}^k$  as the graph with vertex set  $[k]$  where  $V_i$  and  $V_j$  forms an edge if and only if the pair  $(V_i, V_j)$  is  $\varepsilon$ -regular. Note that condition (iii) is equivalent to saying that the  $\varepsilon$ -reduced graph of the partition has minimum degree at least  $(1 - \varepsilon)k$ . For a real number  $\delta$ , we define the  $(\varepsilon, \delta)$ -reduced graph of a partition  $\{V_i\}_{i=0}^k$  as the graph with vertex set  $[k]$  where  $V_i$  and  $V_j$  forms an edge if and only if the pair  $(V_i, V_j)$  is  $\varepsilon$ -regular with density at least  $\delta$ . The celebrated regularity lemma asserts that all large graphs admit an  $\varepsilon$ -regular partition (see [21] for the version of the regularity lemma as stated here).

**Theorem 2.1.** *For all  $\varepsilon$  and  $t$ , there exists  $n_0 = n_0(\varepsilon, t)$  and  $T = T(\varepsilon, t)$  such that the following holds for all  $n \geq n_0$ . Every  $n$ -vertex graph  $G$  admits an  $\varepsilon$ -regular partition into  $k$  parts where  $t \leq k \leq T$ .*

We will later need an  $\varepsilon$ -regular partition with a prescribed number of parts. Such partition can be produced by taking a random refinement of an  $\varepsilon$ -regular partition obtained through the regularity lemma. The following lemma, proved in [14], can be used to verify that such partition indeed works. It asserts that a typical pair of subsets of a regular pair is regular.

**Lemma 2.2.** *For  $0 < \beta, \varepsilon < 1$ , there exists  $\varepsilon_0 = \varepsilon_0(\beta, \varepsilon)$  and  $C = C(\varepsilon)$  such that for all  $\varepsilon' \leq \varepsilon_0$  and  $\delta$ , every  $\varepsilon'$ -regular pair  $(X, Y)$  of density at least  $\delta$  satisfies that, for every  $q \geq C\delta^{-1}$ , the number of sets  $Q \subseteq X$  of cardinality  $q$  that form an  $\varepsilon$ -regular pair of density at least  $\delta$  with  $Y$  is at least  $(1 - \beta^q) \binom{|V_1|}{q}$ .*

By combining Theorem 2.1 and Lemma 2.2, we can prove a regularity lemma which outputs a partition with a prescribed number of parts.

**Lemma 2.3.** *For all  $\varepsilon$ , there exists  $T = T(\varepsilon)$  such that for all  $k \geq T$  there exists  $n_0(\varepsilon, k)$  such that the following holds for all  $n \geq n_0$ . Every  $n$ -vertex graph  $G$  admits an  $\varepsilon$ -regular partition  $V_0 \cup V_1 \cup \dots \cup V_k$ .*

*Proof.* Let  $\varepsilon_0 = \min\{\varepsilon_0\}_{2.2}(\frac{1}{2}, \varepsilon), \frac{\varepsilon}{2}\}$ ,  $C = C_{2.2}(\varepsilon)$ , and  $T = \frac{2}{\varepsilon}T_{2.1}(\varepsilon_0, \frac{1}{\varepsilon_0})$ . Suppose that an integer  $k \geq T$  is given. Apply Theorem 2.1 with  $\varepsilon_{2.1} = \varepsilon_0$  and  $t_{2.1} = \frac{1}{\varepsilon_0}$  to find an  $\varepsilon_0$ -regular partition  $V_0 \cup V_1 \cup \dots \cup V_r$  where  $\frac{1}{\varepsilon_0} \leq r \leq \frac{\varepsilon}{2}T$ . Define  $s = \lceil \frac{k}{r} \rceil$  and note that  $s \geq \frac{T}{r} \geq \frac{2}{\varepsilon}$ .

For each  $i \in [r]$ , we may assume that  $|V_i|$  is divisible by  $s$  by moving at most  $s - 1$  vertices from  $V_i$  to  $V_0$  if necessary. For each  $i \in [r]$ , let  $V_i = V_{i,1} \cup V_{i,2} \cup \dots \cup V_{i,s}$  be a partition chosen uniformly at random where  $|V_{i,j}| = \frac{1}{s}|V_i|$  for all  $j \in [s]$ . By Lemma 2.2, if  $(V_i, V_{i'})$  is  $\varepsilon_0$ -regular, then for all  $j, j' \in [s]$ , the probability that  $(V_{i,j}, V_{i',j'})$  forms an  $\varepsilon$ -regular pair is at least  $1 - 2^{-\Omega(n)}$ . Hence by

---

<sup>1</sup>We deviate from the standard practice enforcing at most  $\varepsilon k^2$  pairs that are not  $\varepsilon$ -regular.

the union bound, we can find partitions  $V_i = V_{i,1} \cup V_{i,2} \cup \dots \cup V_{i,s}$  for each  $i \in [r]$  so that  $(V_{i,j}, V_{i',j'})$  forms an  $\varepsilon$ -regular pair whenever  $(V_i, V_{i'})$  forms an  $\varepsilon_0$ -regular pair. Thus each  $V_{i,j}$  forms an  $\varepsilon$ -regular pair with at least  $(1 - \varepsilon_0)rs$  other sets  $V_{i',j'}$ . Arbitrarily remove  $rs - k \leq r - 1$  parts  $(i, j) \in [r] \times [s]$ , combine the removed sets with  $V_0$  and re-label the sets so that we obtain a partition  $U_0 \cup U_1 \cup \dots \cup U_k$  of the vertex set. For each  $i \in [k]$ , there are at most  $\varepsilon_0 rs \leq \frac{\varepsilon}{2}rs \leq \varepsilon k$  other indices  $i' \in [k]$  for which  $(U_i, U_{i'})$  is not  $\varepsilon$ -regular. Moreover,  $|U_0| \leq \varepsilon_0 n + (s-1)r + (r-1)\frac{n}{sr} \leq \varepsilon n$  and therefore we found a partition with the desired properties.  $\square$

The blow-up lemma, developed by Komlós, Sárközy, and Szemerédi [20], is a powerful tool used in embedding large subgraphs. Informally, quoting Komlós, Sárközy, and Szemerédi, it asserts that, “regular pairs behave like complete bipartite graphs from the point of view of bounded degree subgraphs.”. We use the following version of the blow-up lemma.

**Lemma 2.4.** *For all  $\xi, \delta, \Delta$ , there exists  $\varepsilon = \varepsilon(\xi, \delta, \Delta)$  such that the following holds for all natural numbers  $k$  if  $m \geq m_0$  for some sufficiently large  $m_0 = m_0(k, \varepsilon)$ . Let  $G$  be a graph with maximum degree at most  $\Delta$ . Let  $\Gamma$  be a graph with a vertex partition  $\{V_i\}_{i=1}^k$  satisfying  $|V_i| \geq (1 + \xi)m$  for all  $i \in [k]$ , and let  $R$  be its  $(\varepsilon, \delta)$ -reduced graph. Suppose that there exists a homomorphism  $f$  from  $G$  to  $R$  where  $|f^{-1}(i)| \leq m$  for all  $i \in [k]$ . Then there exists an embedding of  $G$  to  $\Gamma$ .*

Lemma 2.4 differs from the original version of the blow-up lemma in that the restriction on  $\varepsilon$  does not depend on  $R$ . It is a subtle but crucial difference. The proof of Lemma 2.4 is rather technical and hence to avoid unnecessary distraction, we provide it in Section 4.

Theorem 1.4 follows from Lemmas 2.3 and 2.4. We re-state the theorem here.

**Theorem.** *For all  $\Delta, \xi$  and  $\varepsilon$ , there exists  $\beta$  and  $n_0$  such that the following holds for all  $n \geq n_0$ . Let  $G$  and  $H$  be graphs where  $G$  has  $n$  vertices and maximum degree at most  $\Delta$ . Suppose that there exists a homomorphism  $f$  from  $G$  to  $H$  for which  $|f^{-1}(v)| \leq \beta n$  for all  $v \in V(H)$ . Then for the weight-function of  $H$  defined by  $w(v) = \frac{1}{\beta n}|f^{-1}(v)|$  we have  $r(G) \leq (1 + \xi)\hat{r}_\varepsilon(H, w) \cdot \beta n$ .*

*Proof.* By reducing  $\varepsilon$  if necessary, we may assume that  $\varepsilon \leq \min\left\{\varepsilon_{2.4}\left(\frac{\xi}{2}, \frac{1}{2}, \Delta\right), \frac{\xi}{2(1+\xi)}\right\}$ . By the theorem of Conlon, Fox, and Sudakov mentioned in the introduction, there exists a constant  $c$  for which  $r(G) \leq c^{\Delta \log \Delta} n$ . Define  $\beta = T_{2.3}(\varepsilon)^{-1}$  and  $k = \hat{r}_\varepsilon(H)$ . Define  $n_0 = \max\{(n_0)_{2.3}(\varepsilon, \beta^{-1}c^{\Delta \log \Delta}), (m_0)_{2.4}(k, \varepsilon)\}$ . By definition we have  $k = \hat{r}_\varepsilon(H) \geq w(V(H)) = \frac{1}{\beta n}|V(G)| = \beta^{-1}$ . Furthermore since  $r(G) \leq c^{\Delta \log \Delta} n$ , the conclusion holds if  $\hat{r}_\varepsilon(H) \geq \beta^{-1}c^{\Delta \log \Delta}$ . Thus we may assume that  $k = \hat{r}_\varepsilon(H) < \beta^{-1}c^{\Delta \log \Delta}$ . Thus  $\beta^{-1} \leq k < \beta^{-1}c^{\Delta \log \Delta}$ .

Let  $N = (1 + \xi)\hat{r}_\varepsilon(H) \cdot \beta n$ . Suppose that we are given a red/blue coloring of  $K_N$ . Let  $\Gamma_r$  and  $\Gamma_b$  be the red graph and blue graph, respectively. By Theorem 2.1, there exists an  $\varepsilon$ -regular partition  $V_0 \cup V_1 \cup \dots \cup V_k$  of  $\Gamma_r$  (note that it also is an  $\varepsilon$ -regular partition of  $\Gamma_b$ ). Consider the  $\varepsilon$ -reduced graph  $R$  of the partition, and color the edges with red and blue so that an edge  $\{i, j\}$  is red if the red edge density of the pair  $(V_i, V_j)$  is at least  $\frac{1}{2}$  and blue otherwise. Since  $R$  has minimum degree at least  $(1 - \varepsilon)k$ , by the definition of  $\hat{r}_\varepsilon(H)$ , there exists a homomorphism  $g$  from  $H$  to the red subgraph of  $R$  (or the blue subgraph of  $R$ ) such that  $|g^{-1}(i)| \leq 1$ , and thus  $|f^{-1}(g^{-1}(i))| \leq \beta n$

for all  $i \in [k]$ . Without loss of generality, assume that it is to the red subgraph of  $R$ . Note that  $h := g \circ f$  is a homomorphism from  $G$  to the red subgraph of  $R$  satisfying  $|h^{-1}(i)| \leq \beta n$  for all  $i \in [k]$ . Further note that for each  $i \in [k]$ , we have  $|V_i| \geq \frac{1-\varepsilon}{k}N \geq (1 + \frac{\xi}{2})n$ . Therefore by Lemma 2.4, we can find a copy of  $G$  in  $\Gamma_r$ .  $\square$

### 3 Weighted Ramsey number

#### 3.1 Wheel graph

Before proving Theorem 1.3, we first show that the weighted Ramsey number of wheel graphs is small as claimed in the introduction without proof. Recall that a wheel graph  $W_k$  is a graph with  $k$  vertices consisting of a cycle on  $k - 1$  vertices and a vertex adjacent to all vertices on the cycle. Let  $w : V(W_k) \rightarrow [0, 1]$  be a weight function and define  $m = w(V(W_k))$  as the total weight. Let  $C$  be the cycle obtained from  $W_k$  by removing the vertex of degree  $k - 1$ . By restricting the domain of  $w$ , we may assume that  $(C, w)$  is a weighted graph. Since a cycle has maximum degree 2, Theorem 1.3 that we will prove in the next subsection implies that there exist constants  $\varepsilon$  and  $c$  such that  $\hat{r}_\varepsilon(C, w) \leq c \cdot w(V(C))$  (where  $c$  is a constant not depending on the order of  $C$ ). By increasing  $c$  and decreasing  $\varepsilon$  if necessary, we may assume that  $c \geq 2$  and  $\varepsilon < \frac{c}{8}$ .

For simplicity, we assume that  $cm$  is an integer. For  $N = 8cm$ , consider a red/blue edge coloring of  $\Gamma$ , where  $\Gamma$  is a graph with  $N$  vertices and minimum degree at least  $(1 - \frac{\varepsilon}{8})N$ . Without loss of generality, we may assume that there exists a vertex  $v_1$  of red degree at least  $\lceil \frac{N-1}{2} \rceil \geq 4cm$ . Let  $X$  be an arbitrary set of red neighbors of  $v_1$  of size exactly  $4cm$  and let  $\Gamma_1$  be the graph induced on  $X$ . If there exists a vertex  $v_2$  of blue degree at least  $cm$  in  $\Gamma_1$ , then let  $Y$  be an arbitrary set of blue neighbors of  $v_2$  in  $\Gamma_1$  of size exactly  $cm$  and let  $\Gamma_2$  be the subgraph of  $\Gamma_1$  induced on  $Y$ . Note that  $\Gamma_2$  has minimum degree at least  $cm - \frac{\varepsilon}{8}N = (1 - \varepsilon)cm$ . Therefore, we can find a monochromatic copy of  $(C, w)$  in  $\Gamma_2$ . If it is red, then together with  $v_1$ , it forms a monochromatic copy of  $(W_k, w)$ , and if it is blue, then together with  $v_2$ , it forms a monochromatic copy of  $(W_k, w)$ .

Hence we may assume that all vertices of  $\Gamma_1$  has blue degree at most  $cm - 1$  in  $\Gamma_1$ . Since  $\Gamma$  has minimum degree at least  $(1 - \varepsilon)N$ , it follows that  $\Gamma_1$  has minimum degree at least  $4cm - \varepsilon N$ . Therefore  $\Gamma_1$  has minimum red degree at least  $3cm - \varepsilon N$ . Let the vertices of  $C$  be  $x_1, x_2, \dots, x_{k-1}$  in decreasing order of weight (where ties are broken arbitrarily). We will greedily embed the vertices of  $C$  according to this order. Suppose that we finished embedding  $x_1, \dots, x_{i-1}$  and let  $\phi$  denote the partial embedding. Note that  $x_i$  has at most two neighbors in  $x_1, \dots, x_{i-1}$ . Suppose that it has two neighbors, and let  $v, v'$  be the images of these vertices in  $\Gamma_1$ . Let  $R$  be the set of common red neighbors of  $v$  and  $v'$ . By the minimum degree condition of  $\Gamma_1$ , we know that  $|R| \geq 2cm - 2\varepsilon N$ . If there exists a vertex  $v'' \in R$  such that  $w(\phi^{-1}(v'')) + w(x_i) \leq 1$ , then we define  $\phi(x_i) = v''$ . Otherwise, we have  $w(\phi^{-1}(v'')) > 1 - w(x_i) \geq 0$  for all  $v'' \in R$ . If  $w(x_i) \leq \frac{1}{2}$ , then it implies that  $w(\phi^{-1}(v'')) > \frac{1}{2}$  for all  $v'' \in R$ . On the other hand, if  $w(x_i) > \frac{1}{2}$ , then we have  $w(x_j) > \frac{1}{2}$  for all  $j < i$ . Therefore  $w(\phi^{-1}(v'')) > 0$  implies that  $\phi^{-1}(v'') \neq \emptyset$ , and thus  $w(\phi^{-1}(v'')) > \frac{1}{2}$  for all  $v'' \in R$ . In both cases, we have  $w(\phi^{-1}(v'')) > \frac{1}{2}$  for all  $v'' \in R$ . Hence

$w(\phi^{-1}(R)) > \frac{1}{2}|R| \geq cm - \varepsilon N > m$  which contradicts the fact that  $m = w(V(W_k))$ . Therefore we can define  $\phi(x_i)$  as above and continue the process. The other case when there are less than two neighbors of  $x_i$  in  $x_1, \dots, x_{i-1}$  can be similarly handled.

### 3.2 Bounded degree graphs

In this section, we adapt the proof of Conlon, Fox, and Sudakov [9] to prove Theorem 1.3. We say that a graph  $\Gamma$  is *bi- $(\varepsilon, \delta)$ -dense* if for all disjoint pairs of vertex subsets  $X, Y \subseteq V(\Gamma)$  of sizes at least  $|X|, |Y| \geq \varepsilon|V(\Gamma)|$ , we have  $d(X, Y) \geq \delta$ . The following definition is essentially from [9] (we added an additional parameter  $\delta$ ).

**Definition 3.1.** A graph  $\Gamma$  on  $N$  vertices is  $(\alpha, \beta, \rho, \delta, \Delta)$ -dense if there is a sequence  $U_1, U_2, \dots, U_s$  of disjoint vertex subsets each of cardinality at least  $\alpha N$  and non-negative integers  $d_1, \dots, d_s$  such that  $d_1 + \dots + d_s = \Delta - s + 1$ , and the following holds:

- (i) For all  $i \in [s]$ , the induced subgraph  $\Gamma[U_i]$  is bi- $(\rho^{2d_i}, \delta)$ -dense, and
- (ii) for  $1 \leq i < j \leq s$ , each vertex in  $U_i$  has at least  $(1 - \beta)|U_j|$  neighbors in  $U_j$ .

Note that monotonicity holds in a sense that if a graph is  $(\alpha', \beta', \rho', \delta', \Delta')$ -dense and  $\alpha' \geq \alpha, \beta' \leq \beta, \rho' \leq \rho, \delta' \geq \delta, \Delta' \geq \Delta$ , then it is also  $(\alpha, \beta, \rho, \delta, \Delta)$ -dense. The following lemma was proved in [9, Lemma 2.2].

**Lemma 3.2.** Let  $D = 2^h - 1$  for a non-negative integer  $h$ , and  $\rho$  be a fixed real number. For all  $N \geq 1$ , every edge-coloring of  $K_N$  with two colors red and blue, the red graph or the blue graph is  $(2^{-2h}\rho^{6D-4h}, 2(D+1)\rho, \rho, \rho, D)$ -dense.

The ‘stable version’ of the lemma above immediately follows for small values of  $\varepsilon$ .

**Lemma 3.3.** Let  $D = 2^h - 1$  for a non-negative integer  $h$ , and  $\rho, \varepsilon$  be positive real numbers satisfying  $\varepsilon < 2^{-2h-1}\rho^{8D-4h+1}$ . If  $\Gamma$  is a graph on  $N$  vertices with minimum degree at least  $(1 - \varepsilon)N$ , then for every edge-coloring of  $\Gamma$  with two colors red and blue, the red graph or the blue graph is  $(2^{-2h}\rho^{6D-4h}, 4(D+1)\rho, \rho, \frac{1}{2}\rho, D)$ -dense.

*Proof.* Define  $\alpha = 2^{-2h}\rho^{6D-4h}$  and  $\beta = 2(D+1)\rho$ . Consider an edge-coloring of  $K_N$  with three colors, where an edge has color red (blue) if it is an edge of color red (blue) in  $\Gamma$ , and has color green if it is not an edge in  $\Gamma$ . Let  $\Gamma_r, \Gamma_b, \Gamma_g$  be the graph consisting of red, blue, and green edges respectively. By Lemma 3.2, either  $\Gamma_r \cup \Gamma_g$  or  $\Gamma_b$  is  $(\alpha, \beta, \rho, \rho, D)$ -dense. If  $\Gamma_b$  is  $(\alpha, \beta, \rho, \rho, D)$ -dense, then the conclusion immediately follows by monotonicity.

Hence we may assume that  $\Gamma' = \Gamma_r \cup \Gamma_g$  is  $(\alpha, \beta, \rho, \rho, D)$ -dense. By definition, there exists a sequence  $U_1, U_2, \dots, U_s$  of disjoint vertex subsets each of cardinality at least  $\alpha N$  and non-negative integers  $d_1, \dots, d_s$  such that  $d_1 + \dots + d_s = D - s + 1$  where the following holds:

- (i) For all  $i \in [s]$ , the induced subgraph  $\Gamma'[U_i]$  is bi- $(\rho^{2d_i}, \rho)$ -dense, and
- (ii) for  $1 \leq i < j \leq s$ , each vertex in  $U_i$  has at least  $(1 - \beta)|U_j|$  neighbors in  $U_j$  in  $\Gamma'$ .

We claim that  $\Gamma_r[U_i]$  is bi- $(\rho^{2d_i}, \frac{1}{2}\rho)$ -dense for all  $i \in [s]$ , and that for  $1 \leq i < j \leq s$ , each vertex in  $U_i$  has at least  $(1 - 2\beta)|U_j|$  neighbors in  $U_j$  in  $\Gamma_r$ . Note that this proves the lemma. To prove the first part of the claim, fix an index  $i \in [s]$  and consider a pair of disjoint vertex subsets  $X, Y \subseteq U_i$  of sizes  $|X|, |Y| \geq \rho^{2d_i}|U_i| \geq \rho^{2D}\alpha N = 2^{-2h}\rho^{8D-4h}$ . By Property (i), we have  $e_{\Gamma'}(X, Y) \geq \rho|X||Y|$ . Therefore since  $\varepsilon N \leq \frac{1}{2}\rho|Y|$ ,

$$e_{\Gamma_r}(X, Y) \geq e_{\Gamma'}(X, Y) - |X| \cdot \varepsilon N \geq \rho|X||Y| - \frac{1}{2}\rho|X||Y| = \frac{1}{2}\rho|X||Y|.$$

To prove the second part of the claim, fix two indices  $i, j \in [s]$  satisfying  $i < j$ . By Property (ii), each vertex  $u \in U_i$  has at least  $(1 - \beta)|U_j|$  neighbors in  $U_j$  in  $\Gamma'$ . Since  $\varepsilon \leq \alpha\beta$ , we see that  $u$  has at least  $(1 - \beta)|U_j| - \varepsilon N \geq (1 - 2\beta)|U_j|$  neighbors in  $U_j$ , thus proving the claim.  $\square$

We also need the following theorem proved by Lovász [23].

**Lemma 3.4.** *Let  $G$  be a graph of maximum degree at most  $\Delta$ , and  $d_1, \dots, d_s$  be non-negative integers satisfying  $d_1 + \dots + d_s \geq \Delta - t + 1$ . Then there exists a vertex partition  $V(G) = V_1 \cup \dots \cup V_s$  such that for all  $i \in [s]$ , the induced subgraph  $G[V_i]$  has maximum degree at most  $d_i$ .*

The following lemma is an embedding lemma for  $(\alpha, \beta, \rho, \delta, \Delta)$ -dense graphs. It is a variant of [9, Lemma 2.5] for weighted graphs.

**Lemma 3.5.** *Let  $\alpha, \rho, \delta$  be fixed positive real numbers satisfying  $\rho \leq \frac{1}{16}$  and  $\delta \geq \frac{\rho}{2}$ . If  $\Gamma$  is an  $(\alpha, \frac{1}{2\Delta}, \rho, \delta, \Delta)$ -dense graph on  $N \geq 8\alpha^{-1}(\frac{2}{\delta})^\Delta n$  vertices, then  $\Gamma$  contains a copy of every weighted graph  $(G, w)$  of total weight at most  $n$  and maximum degree at most  $\Delta$ .*

*Proof.* Note that if  $\Delta = 0$ , then the conclusion trivially holds, and hence we may assume that  $\Delta \geq 1$ . By definition, there exists a sequence  $U_1, U_2, \dots, U_s$  of disjoint vertex subsets of  $\Gamma$  each of cardinality at least  $\alpha N$  and non-negative integers  $\Delta_1, \dots, \Delta_s$  such that  $\Delta_1 + \dots + \Delta_s = \Delta - s + 1$  for which

- (i) the induced subgraph  $\Gamma[U_j]$  is bi- $(\rho^{2\Delta_j}, \delta)$ -dense for each  $j \in [s]$ , and
- (ii) for  $1 \leq j < j' \leq s$ , each vertex in  $U_j$  has at least  $(1 - \frac{1}{2\Delta})|U_{j'}|$  neighbors in  $U_{j'}$ .

Let  $(G, w)$  be a weighted graph of total weight at most  $n$  and maximum degree at most  $\Delta$ . By Lemma 3.4, there exists a vertex partition  $V(G) = V_1 \cup \dots \cup V_s$  such that for all  $j \in [s]$ , the induced subgraph  $G[V_j]$  has maximum degree at most  $\Delta_j$ . Let  $v_1, v_2, \dots, v_n$  be an enumeration of the vertices of  $G$  with the following properties:

- (a) For all  $1 \leq j < j' \leq s$ , the vertices in  $V_j$  come before vertices in  $V_{j'}$ , and
- (b) for all  $j \in [s]$ , the vertices in  $V_j$  are ordered so that their weights form a non-increasing sequence.

For each  $t \in [n]$ , define  $\pi(t) \in [s]$  as the index for which  $v_t \in V_{\pi(t)}$ .

Consider the following greedy algorithm of embedding the vertices of  $G$ , where the  $t$ -th step of the algorithm selects the image of  $v_t$  in  $V(\Gamma)$ . At time  $t$ , the algorithm is given as input a

partial embedding  $f$  defined on  $\{v_1, \dots, v_{t-1}\}$ , where at the initial step,  $f$  is partial embedding of the empty graph. For  $t \in [n]$  and  $i \geq t$ , define  $N_i^{(t)} = N(v_i) \cap \{v_1, \dots, v_{t-1}\}$  as the set of neighbors of  $v_i$  that precede  $v_t$ . Define  $W_i^{(t)} = U_{\pi(i)} \cap \bigcap_{v \in N_i^{(t)}} N(f(v))$  and note that each vertex in  $W_i^{(t)}$  can be used as the image of  $v_t$  to extend the partial embedding. For each  $i \geq t$ , define  $d_i^{(t)} = |N_i^{(t)} \cap V_{\pi(i)}|$  as the number of neighbors of  $v_i$  preceding itself in its own part. Throughout the process, we will maintain the following property:

$$\forall i \geq t, \quad |W_i^{(t)}| \geq \frac{1}{2} \left(\frac{\delta}{2}\right)^{d_i^{(t)}} |U_{\pi(i)}|. \quad (1)$$

Note that if  $d_i^{(t)} = 0$ , then (1) follows from Property (ii) since  $|N_i^{(t)}| \leq \deg(v_t) \leq \Delta$  holds and

$$|W_i^{(t)}| \geq |U_{\pi(i)}| - \frac{1}{2\Delta} |U_{\pi(i)}| \cdot |N_i^{(t)}| \geq \frac{1}{2} |U_{\pi(i)}|. \quad (2)$$

Initially at  $t = 1$ , we define  $W_i^{(1)} = U_{\pi(i)}$  for all  $i \in [n]$ . Moreover since  $d_i^{(1)} = 0$  for all  $i \in [n]$ , equation (1) holds.

Suppose that we are at the  $t$ -th step of the algorithm for some  $t \in [n]$ . Define  $I^+ = \{i > t : v_i \in N(v_t) \cap V_{\pi(t)}\}$ . Let  $W \subseteq W_t^{(t)}$  be the set of vertices  $u \in W_t^{(t)}$  such that for all  $i \in I^+$ ,  $|N(u) \cap W_i^{(t)}| \geq \frac{\delta}{2} |W_i^{(t)}|$ . If there exists a vertex  $u \in W$  such that  $w(f^{-1}(u)) \leq 1 - w(v_t)$ , then define  $f(v_t) = u$ . This is a partial embedding since  $f(v_t) \in W_t^{(t)}$  and  $w(f^{-1}(u)) \leq 1$ . Furthermore (1) is satisfied for  $i > t$  having  $\pi(i) > \pi(t)$  by (2), and having  $\pi(i) = \pi(t)$  but  $i \notin I^+$  since  $W_i^{(t+1)} = W_i^{(t)}$ . If  $\pi(i) = \pi(t)$  and  $i \in I^+$ , then  $d_i^{(t+1)} = d_i^{(t)} + 1$  and therefore (1) holds since  $W_i^{(t+1)} = N(u) \cap W_i^{(t)}$ .

Therefore it suffices to prove the existence of a vertex  $u \in W$  satisfying  $w(f^{-1}(u)) \leq 1 - w(v_t)$ . Suppose that all vertices  $u \in W$  satisfy  $w(f^{-1}(u)) > 1 - w(v_t)$ . Recall that  $w(v_j) \geq w(v_t)$  for all  $j \leq t$  satisfying  $\pi(j) = \pi(t)$  by Property (b). Since  $w(f^{-1}(u)) > 1 - w(v_t) \geq 0$  implies that  $f^{-1}(u) \neq \emptyset$ , if  $w(v_t) > \frac{1}{2}$ , then it follows that  $w(f^{-1}(u)) \geq w(v_t) > \frac{1}{2}$ . On the other hand, if  $w(v_t) \leq \frac{1}{2}$ , then  $w(f^{-1}(u)) > 1 - w(v_t) \geq \frac{1}{2}$ . Therefore for all vertices  $u \in W$ , we have  $w(f^{-1}(u)) > \frac{1}{2}$ . Since

$$\frac{1}{2} |W| < \sum_{u \in W} w(f^{-1}(u)) \leq w(V(G)) \leq n,$$

we see that  $|W| < 2n$ .

For all vertices  $u \in W_t^{(t)} \setminus W$ , there exists  $i \in I^+$  such that  $|N(u) \cap W_i^{(t)}| < \frac{\delta}{2} |W_i^{(t)}|$ . For notational simplicity, define  $k = \Delta_{\pi(t)}$ . Since  $|I_+| \leq k$ , by the pigeonhole principle, there exists an index  $i_0 \in I^+$  such that  $|N(u) \cap W_{i_0}^{(t)}| < \frac{\delta}{2} |W_{i_0}^{(t)}|$  holds for at least  $\frac{1}{k} |W_t^{(t)} \setminus W|$  vertices. Let  $X_1$  be the set of these vertices and note that

$$|X_1| \geq \frac{1}{k} (|W_t^{(t)}| - 2n) \geq \frac{1}{k} \left( \frac{1}{2} \left(\frac{\delta}{2}\right)^k |U_{\pi(t)}| - 2n \right) \geq \frac{1}{4k} \left(\frac{\delta}{2}\right)^k |U_{\pi(t)}| \geq \rho^{2k} |U_{\pi(t)}|,$$

where the second to last inequality follows since  $|U_{\pi(t)}| \geq \alpha N$  and the last inequality follows since  $\rho \leq \frac{1}{16}$  and  $\rho \leq 2\delta$ .

Define  $X_2 = W_{i_0}^{(t)}$  and note that (1) implies  $|X_2| \geq \frac{1}{2}(\frac{\delta}{2})^k |U_{\pi(t)}| \geq 2\rho^{2k} |U_{\pi(t)}|$ . Let  $X'_1 \subseteq X_1$  be an arbitrary subset of size exactly  $\rho^{2k} |U_{\pi(t)}|$ , and define  $X'_2 = X_2 \setminus X'_1$ . Then  $|X'_2| \geq \frac{1}{2}|X_2| \geq \rho^{2k} |U_{\pi(t)}|$ . Furthermore, each vertex  $w \in X'_1$  has at most  $\frac{\delta}{2}|X_2| \leq \delta|X'_2|$  neighbors in  $X'_2$ . This contradicts the fact that  $\Gamma[U_{\pi(t)}]$  is bi- $(\rho^{2k}, \delta)$ -dense. Therefore there exists a vertex  $u \in W$  satisfying  $w(f^{-1}(u)) \leq 1 - w(v_t)$ .  $\square$

Theorem 1.3 straightforwardly follows from Lemmas 3.3 and 3.5.

**Theorem.** *There exists a constant  $c > 1$  such that the following holds for all  $\Delta$  and  $\varepsilon$  satisfying  $\varepsilon < c^{-\Delta \log \Delta}$ . If  $(G, w)$  is a weighted graph with maximum degree at most  $\Delta$  and total weight at most  $n$ , then  $\hat{r}_\varepsilon(G) \leq c^{\Delta \log \Delta} n$ .*

*Proof.* Let  $N = c^{\Delta \log \Delta} n$  for a constant  $c$  to be chosen later. Let  $(G, w)$  be a weighted graph given as above. Suppose that  $\Gamma$  is a graph on  $N$  vertices with minimum degree at least  $(1 - \varepsilon)N$ , and consider an edge-coloring with two colors red and blue. Let  $h$  be the integer satisfying  $2^{h-1} \leq \Delta \leq 2^h - 1$  and note that  $h \leq \log(2\Delta)$ . Define  $D = 2^h - 1 \leq 2\Delta$  and  $\rho = \frac{1}{2^{h+4}\Delta} \geq \frac{1}{32\Delta^2}$ . If  $c$  is sufficiently large, then  $\varepsilon < 2^{-2h-1}\rho^{8D-4h+1}$  and thus by Lemma 3.3, we see that the red graph or the blue graph is  $(2^{-2h}\rho^{6D-4h}, 4(D+1)\rho, \rho, \frac{1}{2}\rho, D)$ -dense. Without loss of generality, assume that it is the red graph. Then by monotonicity, the red graph is  $(\rho^{12D}, \frac{1}{2\Delta}, \rho, \frac{1}{2}\rho, \Delta)$ -dense. Therefore if  $c$  is large enough, then  $N \geq 8\rho^{-12D}(\frac{4}{\rho})^\Delta n$  and by Lemma 3.5, the red graph contains a copy of  $G$ .  $\square$

## 4 A variant of the blow-up lemma

In this section, we prove Lemma 2.4, a variant of the blow-up lemma. We will use a simplified version of the Random Greedy Algorithm (RGA) developed by Komlós, Sárközy, and Szemerédi [20]. Their original algorithm consisted of two phases. In Phase 1, they embed the vertices one at a time, where at each step one considers all possible images that is consistent with the previous embedding and choose a random vertex among them. Phase 1 continues until almost all vertices of the graph has been embedded. In Phase 2, they finish the embedding by invoking Hall's theorem. For our proof, we do not need the second phase, since we only need an almost spanning embedding. One can prove Lemma 2.4 by carefully making this adjustment in their proof. It is rather straightforward to incorporate this change, but we include the proof here for completeness.

Let  $G$  be a graph with maximum degree at most  $\Delta$ . Let  $\Gamma$  be a graph with a vertex partition  $\{V_i\}_{i=1}^k$  satisfying  $|V_i| \geq (1 + \xi)m$  for all  $i \in [k]$ , and let  $R$  be its  $(\varepsilon, \delta)$ -reduced graph. Suppose that there exists a homomorphism  $f$  from  $G$  to  $R$  where for all  $i \in [k]$ ,  $|f^{-1}(i)| \leq m$  for all  $i \in [k]$ . For simplicity we assume that  $|f^{-1}(i)| = m$  for all  $i \in [k]$  by adding isolated vertices if necessary. In order to avoid confusion, we will refer to the vertices in  $G$  using  $x, y$  and the vertices in  $\Gamma$  using  $v, w$ .

Let  $\varepsilon, \varepsilon_1, \varepsilon_2$  be positive real numbers satisfying  $\varepsilon \ll \varepsilon_2 \ll \varepsilon_1$  where  $\varepsilon_1$  is small enough depending on  $\delta$  and  $\xi$ . We first embed  $f^{-1}(1)$  to  $V_1$ , then  $f^{-1}(2)$  to  $V_2$ , and continue until we embed  $f^{-1}(k)$  to  $V_k$ . Suppose that we finished embedding  $f^{-1}(i-1)$  to  $V_{i-1}$  for some  $i \in [k]$ . Define  $A_0 = f^{-1}(1) \cup \dots \cup f^{-1}(i-1)$  and  $B_0 = V(G) \setminus A_0$ . For each  $y \in B_0$ , define  $N_0(y) = N(y) \cap A_0$  as the set of neighbors of  $y$  already embedded, and let  $d_0(Y) = |N_0(y)|$ . Define  $U_0(y) = V_{f(y)} \cap \bigcap_{z \in N_0(y)} N(\phi(z))$ . Consider the following property:

$\mathcal{P}(i)$  : For all  $X \subseteq V_i$  of size  $\varepsilon_1|V_i| \leq |X| \leq m$ , there are less than  $\varepsilon_1 m$  vertices  $y \in f^{-1}(i)$  such that  $|U_0(y) \cap X| \geq (1 - \varepsilon_1)|U_0(y)|$ .

We will show that there exists a random embedding algorithm that embeds  $f^{-1}(i)$  to  $V_i$  so that the probability that  $\mathcal{P}(1), \dots, \mathcal{P}(i)$  hold but  $\mathcal{P}(i+1)$  does not is small.

Fix an arbitrary enumeration of the vertices in  $f^{-1}(i)$ . We will iteratively embed the vertices of  $f^{-1}(i)$  mainly following the order of this enumeration. For some  $s \geq 0$ , suppose that we finished embedding  $s$  vertices of  $f^{-1}(i)$  and let  $A_s \subseteq V(G)$  be the set of embedded vertices and  $B_s = V(G) \setminus A_s$  be its complement. Hence  $|A_s \setminus A_0| = s$ . Let  $\phi$  be the partial embedding of  $G$  to  $\Gamma$  defined on  $A_s$ . We will maintain a first-in first-out queue  $Q$  throughout the process, where initially  $Q = \emptyset$ . At the next step, if  $Q \neq \emptyset$ , then we let  $x_s$  be the first vertex in  $Q$ , and if  $Q = \emptyset$ , then we let  $x_s$  be the first non-embedded vertex according to the enumeration given above. We will define the image of  $x_s$  in the next step.

For each vertex  $y \in B_s$ , define  $N_s(y) = N(y) \cap A_s$  as the set of neighbors of  $y$  already embedded, and let  $d_s(Y) = |N_s(y)|$ . Define  $U_s(y) = V_{f(y)} \cap \bigcap_{z \in N_s(y)} N(\phi(z))$ . Throughout the process, we will maintain the following properties:

- (i) for all  $y \in B_s$ , we have  $|U_s(y)| \geq (\delta - \varepsilon)^{d_s(y)}|V_{f(y)}|$ ,
- (ii) for all  $y \in f^{-1}(i) \setminus Q$ , we have  $|U_s(y) \setminus \phi(A_s)| \geq \varepsilon_2|V_i|$ , and
- (iii)  $|Q| \leq \varepsilon_1 m$ .

We will add a vertex to  $Q$  when and only when (ii) fails. Thus Property (ii) always holds. We will later show that Property (i) is maintained by how we choose the embedding  $\phi$ . Note that since we are embedding vertices in  $f^{-1}(i)$  to  $V_i$ , the definition of  $Q$  implies  $Q \subseteq f^{-1}(i)$ . Furthermore since  $f^{-1}(i)$  is an independent set, for all  $y \in B_s \cap f^{-1}(i)$ , we have  $N_s(y) = N_0(y)$  and hence  $U_s(y) = U_0(y)$ . Thus  $|U_t(y) \setminus \phi(A_s)|$  is non-increasing in time  $t$ . For  $y \in Q$ , since  $|U_t(y) \setminus \phi(A_s)| < \varepsilon_2|V_i|$  at the time  $t$  that  $y$  was added to  $Q$ , it follows that if  $y \in Q$  at time  $s$ , then  $|U_s(y) \setminus \phi(A_s)| < \varepsilon_2|V_i|$ . For all  $y \in Q$ , since  $N_s(y) = N_0(y)$  and  $d_s(y) = d_0(y)$  for all  $y \in Q$ , by Property (i),

$$\begin{aligned} |U_s(y) \cap \phi(A_s)| &= |U_s(y)| - |U_s(y) \setminus \phi(A_s)| \\ &= \left(1 - \frac{|U_s(y) \setminus \phi(A_s)|}{|U_s(y)|}\right) |U_0(y)| \\ &\geq \left(1 - \frac{\varepsilon_2}{(\delta - \varepsilon)^\Delta}\right) |U_0(y)| \geq (1 - \varepsilon_1) |U_0(y)|. \end{aligned}$$

If  $|V_i \cap \phi(A_s)| = s \geq \varepsilon_1|V_i|$ , then by Property  $\mathcal{P}(i)$  it follows that  $|Q| \leq \varepsilon_1 m$ . On the other hand if  $s < \varepsilon_1|V_i|$ , then for all  $y \in B_s$ , by Property (i) we have  $|U_s(y) \setminus \phi(A_s)| \geq (\delta - \varepsilon)^{d_s(y)}|V_{f(y)}| - \varepsilon_1|V_{f(y)}| > \varepsilon_1|V_{f(y)}|$  and therefore  $y \notin Q$ . This implies that  $Q = \emptyset$  if  $s < \varepsilon_1|V_i|$ . Therefore Property (iii) holds if  $\mathcal{P}(i)$  holds.

Define  $U = U_s(x_s)$ . Note that if  $x_s \notin Q$ , then  $|U \setminus A_s| \geq \varepsilon_2|V_i|$ . On the other hand, suppose that  $x_s \in Q$ . As observed above,  $d_t(x_s)$  is constant for  $t = 0, 1, 2, \dots, m-1$ . Therefore the size of  $U \setminus A_s$  can change by at most one at each step. Since  $Q$  is a first-in first-out queue, by Property (iii), there are at most  $\varepsilon_1 m$  steps between the time that  $x_s$  was first added to the queue and time  $s$ . This implies that  $|U \setminus A_s| \geq \varepsilon_2|V_i| - \varepsilon_1 m \geq \frac{1}{2}\varepsilon_2|V_i|$ . Therefore in both cases, we have  $|U \setminus A_s| \geq \frac{1}{2}\varepsilon_2|V_i|$ .

For each  $y \in N(x_s) \cap B_s$ , since  $f$  is a homomorphism from  $G$  to  $R$ , we know that the pair  $(V_i, V_{f(y)})$  is  $\varepsilon$ -regular of density at least  $\delta$ . Moreover since  $U \subseteq V_i, U_s(y) \subseteq V_{f(y)}$ , and  $|U_s(y)| \geq (\delta - \varepsilon)^{d_s(y)}|V_{f(y)}|$  (by Property (i)), the set of vertices  $Z_y \subseteq U$  with less than  $(\delta - \varepsilon)|U_s(y)|$  neighbors in  $U_s(y)$  has size  $|Z_y| \leq \varepsilon|V_i|$ . Define  $U' = (U \setminus A_s) \setminus \bigcup_{y \in N(x_s) \cap B_s} Z_y$  and note that

$$|U'| \geq \frac{1}{2}\varepsilon_2|V_i| - \Delta\varepsilon|V_i| \geq \frac{1}{4}\varepsilon_2|V_i|.$$

Let  $\phi(x_s)$  be a vertex in  $U'$  chosen uniformly at random. The following lemma shows that Property  $\mathcal{P}(i+1)$  holds with high probability after we finish embedding  $f^{-1}(i)$  to  $V_i$ .

**Lemma 4.1.** *The probability that  $\mathcal{P}(i+1)$  does not hold but  $\mathcal{P}(1), \dots, \mathcal{P}(i)$  holds is at most  $e^{-\Omega(m)}$ .*

Given this lemma, by taking the union bound, we see that the probability that  $\mathcal{P}(i)$  does not hold for some  $i$  is at most  $ke^{-\Omega(m)} = o(1)$ . Hence with non-zero probability, the algorithm will successfully terminate and embed  $G$  to  $\Gamma$ .

*Proof of Lemma 4.1.* Let  $\mathcal{E}$  be the event that  $\mathcal{P}(1), \dots, \mathcal{P}(i)$  holds. Fix a set  $X \subseteq V_{i+1}$  of size  $\varepsilon_1|V_i| \leq |X| \leq m$ . Define  $A = \bigcup_{j=1}^i f^{-1}(j)$  and  $B = V(G) \setminus A$ . For each  $y \in f^{-1}(i+1)$ , define  $U(y) = V_{i+1} \cap \bigcap_{z \in N(y) \cap A} N(\phi(z))$ . Let  $R \subseteq f^{-1}(i+1)$  be a fixed set of size at least  $\varepsilon_1 m$ . We first compute the probability that

$$(*) \text{ all vertices } y \in R \text{ satisfies } |U(y) \cap X| \geq (1 - \varepsilon_1)|U(y)|.$$

Note that  $\mathcal{P}(i+1)$  holds if there are no such pair of sets  $(X, R)$ .

Since  $G$  has maximum degree at most  $\Delta$ , we can find a subset  $R' \subseteq R$  of size at least  $\frac{|R|}{\Delta^2 + 1}$  whose pairwise distance is at least 3 in  $G$ . In other words, the sets  $N(y)$  are disjoint for vertices  $y \in R'$ . Fix a vertex  $y \in R'$ . We examine the probability that  $|U(y) \cap X| \geq (1 - \varepsilon_1)|U(y)|$ . Let  $z_1, z_2, \dots, z_d$  be the vertices in  $A \cap N(y)$  in the order of embedding (note that  $d \leq \Delta$ ). Then

$$U(y) = V_{i+1} \cap N(\phi(z_1)) \cap N(\phi(z_2)) \cdots \cap N(\phi(z_d)).$$

For  $j = 0, 1, 2, \dots, d$ , define  $W_j(y) = V_{i+1} \cap N(\phi(z_i)) \cap \dots \cap N(\phi(z_j))$ . By the definition of our embedding algorithm, either  $\mathcal{E}$  does not hold, or we have  $|W_j(y)| \geq (\delta - \varepsilon)^j |V_{i+1}|$  for all  $j = 1, 2, \dots, d$ . Since  $U(y) = W_d(y)$ , we have

$$|U(y) \cap X| \geq (1 - \varepsilon_1) |U(y)| \geq (1 - \varepsilon_1) \cdot (\delta - \varepsilon)^d |V_{i+1}| > (\delta + \varepsilon)^d |X|.$$

Therefore there exists some  $t$  such that

$$|W_t(y) \cap X| > (\delta + \varepsilon)^t |X| \quad \text{but} \quad |W_{t-1}(y) \cap X| \leq (\delta + \varepsilon)^{t-1} |X|. \quad (3)$$

Since  $|X| \geq \varepsilon_1 |V_{i+1}|$  and  $(\delta + \varepsilon)^{\Delta} \varepsilon_1 \geq \varepsilon$ , the above can hold only if  $|W_t(y) \cap X| \geq \varepsilon |V_{i+1}|$ . This implies that  $|W_{t-1}(y) \cap X| \geq \varepsilon |V_{i+1}|$ . Furthermore, since  $f$  is a homomorphism from  $G$  to  $R$ , the pair  $(V_{i+1}, V_{f(z_t)})$  is  $\varepsilon$ -regular with density at least  $\delta$ . Since  $X \subseteq V_{i+1}$ , there are at most  $\varepsilon |V_{f(z_t)}|$  vertices  $z \in V_{f(z_t)}$  for which defining  $\phi(z_t) = z$  would cause (3). Thus we can conclude that  $y \in R'$  only if there exists  $z_y \in A \cap N(y)$  whose image  $\phi(z_y)$  was chosen in a set of size at most  $\varepsilon |V_{f(z_y)}|$ .

Therefore (\*) holds only if for each  $y \in R'$ , there exists  $z_y \in A \cap N(y)$  as above. On the other hand if  $\mathcal{E}$  holds, then  $\phi(z_y)$  was chosen inside a subset of  $V_{f(z_y)}$  of size at least  $\frac{1}{4} \varepsilon_2 |V_{f(z_y)}|$ . Since the vertices in  $R'$  have pairwise distance at least 3, all these vertices are distinct. Moreover, the number of choices of these vertices  $z_y$  is at most  $\Delta^{|R'|}$  and thus the probability that  $\mathcal{E}$  holds and (\*) holds is at most

$$\Delta^{|R'|} \cdot \left( \frac{\varepsilon}{\varepsilon_2/4} \right)^{|R'|} \leq \left( \frac{4\varepsilon\Delta}{\varepsilon_2} \right)^{|R'|/(\Delta^2+1)} \leq \left( \frac{4\varepsilon\Delta}{\varepsilon_2} \right)^{\varepsilon_1 m / (\Delta^2+1)}.$$

The number of choices for  $R$  is at most  $2^m$ . Since the size of  $X$  satisfies  $\varepsilon_1 |V_i| \leq |X| \leq m$ , we must have  $|V_i| \leq \varepsilon_1^{-1} m$  or otherwise the lemma is vacuously true. Therefore the number of choices for the set  $X$  is at most  $2^{\varepsilon_1^{-1} m}$ . Hence if  $\varepsilon$  is sufficiently small, then the lemma follows from the union bound.  $\square$

## 5 Bipartite graphs of small bandwidth

In this section we prove Theorem 5.2. Our proof is based on a variant on the idea independently used by Conlon [8], and by Fox and Sudakov [18] based on dependent random choice. This variant of dependent random choice has been recently used in [22] to establish some embedding results for degenerate graphs. The following lemma is the main ingredient of the proof.

**Lemma 5.1.** *Let  $G$  be an  $n$ -vertex graph of minimum degree at least  $\alpha n$  and let  $X_0$  be a subset of vertices. For every positive real number  $\beta$ , there exists a set  $X \subseteq V(G)$  satisfying the following properties:*

- (i)  $|X| \geq \frac{1}{2} \alpha^{2\Delta} |V(G)|$ ,
- (ii)  $|X \cap X_0| \geq \frac{1}{2} \alpha^{2\Delta} |X_0|$ , and
- (iii) the number of  $\Delta$ -tuples in  $X^\Delta$  with less than  $\beta n$  common neighbors is at most  $(\frac{2\beta}{\alpha^{2\Delta}} |X|)^\Delta$ .

*Proof.* Define  $V = V(G)$ . Choose  $\Delta$  vertices  $\mathbf{v}_1, \dots, \mathbf{v}_\Delta \in V$  independently and uniformly at random, and let  $\mathbf{X} = \bigcap_{i=1}^\Delta N(\mathbf{v}_i)$ . By linearity of expectation,

$$\begin{aligned}\mathbb{E}[|X_0 \cap \mathbf{X}| \cdot |\mathbf{X}|] &= \sum_{x \in X_0, y \in V} \mathbf{P}(x, y \in \mathbf{X}) = \sum_{x \in X_0, y \in V} \left( \frac{\text{codeg}(x, y)}{n} \right)^\Delta \\ &\geq |X_0|n \left( \frac{1}{|X_0|n^2} \sum_{x \in X_0} \sum_{y \in V} \text{codeg}(x, y) \right)^\Delta,\end{aligned}\quad (4)$$

where the inequality follows from convexity. For a fixed vertex  $x \in X_0$ , the sum  $\sum_{y \in V} \text{codeg}(x, y)$  counts the number of walks of length 2 in  $V$  that starts at  $x$ . Since  $G$  has minimum degree at least  $\alpha n$ , for all  $x \in X_0$ , we have  $\sum_{y \in V} \text{codeg}(x, y) \geq (\alpha n)^2$ . Hence from (4),

$$\mathbb{E}[|X_0 \cap \mathbf{X}| \cdot |\mathbf{X}|] \geq |X_0|n \left( \frac{|X_0| \cdot \alpha^2 n^2}{|X_0|n^2} \right)^\Delta \geq \alpha^{2\Delta} |X_0|n,$$

and by convexity,

$$\mathbb{E}[|X_0 \cap \mathbf{X}|^\Delta \cdot |\mathbf{X}|^\Delta] \geq \alpha^{2\Delta^2} |X_0|^\Delta n^\Delta.$$

Call a  $\Delta$ -tuple of vertices *bad* if it has less than  $\beta n$  common neighbors. For a set  $A$ , define  $\xi(A)$  as the number of bad  $\Delta$ -tuples in  $A^\Delta$ . The probability of a fixed bad  $\Delta$ -tuple  $T$  being in  $\mathbf{X}^\Delta$  is at most  $(\frac{\text{codeg}(T)}{n})^\Delta \leq \beta^\Delta$ . Hence by linearity of expectation,  $\mathbb{E}[\xi(\mathbf{X})] \leq \beta^\Delta \cdot n^\Delta$ . Since

$$\mathbb{E} \left[ \frac{|X_0 \cap \mathbf{X}|^\Delta \cdot |\mathbf{X}|^\Delta}{\mathbb{E}[|X_0 \cap \mathbf{X}|^\Delta \cdot |\mathbf{X}|^\Delta]} - \frac{\xi(\mathbf{X}) \cdot |X_0 \cap \mathbf{X}|^\Delta}{2\mathbb{E}[\xi(\mathbf{X}) \cdot |X_0 \cap \mathbf{X}|^\Delta]} \right] = \frac{1}{2},$$

there exists a set  $X$  for which

$$\frac{|X_0 \cap X|^\Delta \cdot |X|^\Delta}{\mathbb{E}[|X_0 \cap X|^\Delta \cdot |X|^\Delta]} - \frac{\xi(X) \cdot |X_0 \cap X|^\Delta}{2\mathbb{E}[\xi(X) \cdot |X_0 \cap X|^\Delta]} \geq \frac{1}{2}.$$

In particular,

$$|X_0 \cap X|^\Delta \cdot |X|^\Delta \geq \frac{1}{2} \mathbb{E}[|X_0 \cap X|^\Delta \cdot |X|^\Delta] \geq \frac{1}{2} \alpha^{2\Delta^2} |X_0|^\Delta n^\Delta,$$

and since  $|X_0 \cap X| \leq |X_0|$  and  $|X| \leq n$ , this implies that  $|X| \geq \frac{\alpha^{2\Delta}}{2^{1/\Delta}} n$  and  $|X_0 \cap X| \geq \frac{\alpha^{2\Delta}}{2^{1/\Delta}} |X_0|$  thus proving Properties (i) and (ii). Furthermore,

$$\xi(X) \leq |X|^\Delta \frac{2\mathbb{E}[\xi(X) \cdot |X_0 \cap X|^\Delta]}{\mathbb{E}[|X_0 \cap X|^\Delta \cdot |X|^\Delta]} \leq |X|^\Delta \frac{2\beta^\Delta n^\Delta |X_0|^\Delta}{\alpha^{2\Delta^2} |X_0|^\Delta n^\Delta} \leq |X|^\Delta \left( \frac{2\beta}{\alpha^{2\Delta}} \right)^\Delta,$$

and thus Property (iii) holds.  $\square$

We now prove Theorem 1.6 using Lemma 5.1.

**Theorem.** Let  $\delta$  and  $\alpha$  be positive real numbers. Let  $G$  be an  $n$ -vertex graph of minimum degree at least  $(\delta + \alpha)n$ . Then  $G$  contains all bipartite graphs  $H$  on at most  $\delta n$  vertices with maximum degree at most  $\Delta$  and bandwidth at most  $\frac{1}{256\Delta}\alpha^{6\Delta+1}n$ .

*Proof.* Let  $G$  and  $H$  be graphs given as above. Define  $m = |V(H)|$ . Since  $|V(H)| \leq |V(G)|$ , we can always embed the isolated vertices in the end. Thus we may assume for simplicity that  $H$  has no isolated vertex. Let  $V = V(G)$  and let  $A \cup B$  the bipartition of  $H$ . Define  $\beta = \frac{1}{256\Delta}\alpha^{6\Delta+1}$  and label the vertices of  $H$  using  $[m]$  so that  $|i - j| \leq \beta n$  whenever the vertices with labels  $i$  and  $j$  are adjacent.

For  $t \geq 0$ , define  $B_t := [2t\beta n] \cap B$  and define  $A_t$  as the set of vertices  $a \in A$  for which  $N_H(a) \subseteq B_t$ . Note that since  $H$  has bandwidth at most  $\beta n$ , we have  $(A_{t+1} \cup B_{t+1}) \setminus (A_t \cup B_t) \subseteq ((2t-3)\beta n, (2t+1)\beta n]$ . Therefore

$$|(A_{t+1} \cup B_{t+1}) \setminus (A_t \cup B_t)| < 4\beta n \quad (5)$$

for all  $t \geq 0$ . Note that  $A_0 = B_0 = \emptyset$  since  $H$  has no isolated vertex.

We embed  $H$  into  $G$  using an iterative algorithm. Define  $\gamma = 16\beta\alpha^{-2\Delta}$ . As an initialization, apply Lemma 5.1 to  $G$  with  $(X_0)_{5.1} = V$ ,  $\beta_{5.1} = 8\beta$ , and  $\alpha_{5.1} = \alpha$  to obtain a set  $X_0$  (which is the set  $X$  that we obtain by applying the lemma) of size  $|X_0| \geq \frac{1}{2}\alpha^{2\Delta}n$  where the number of  $\Delta$ -tuples in  $X^\Delta$  with less than  $8\beta n$  common neighbors is at most  $(\gamma|X_0|)^\Delta$ . Define  $\phi$  as the trivial partial embedding of  $H$  to  $G$  defined on  $A_0 \cup B_0 = \emptyset$ .

For  $t \geq 0$ , at the  $t$ -th step of the algorithm, we are given as input a set  $X_t$  and a partial embedding  $\phi$  of  $H$  to  $G$  defined on  $A_t \cup B_t$ . Define  $V_t = V \setminus \phi(A_t \cup B_t)$ . We say that a  $\Delta$ -tuple of vertices  $T$  is  $V_t$ -bad if the number of common neighbors of  $T$  in  $V_t$  is less than  $8\beta n$ . Otherwise, we say that  $T$  is  $V_t$ -good. The given input satisfies the following properties:

- (a)  $X_t \subseteq V_{t-1}$ ,
- (b)  $|X_t| \geq \frac{1}{2}\alpha^{2\Delta+1}n$ ,
- (c)  $\phi(B_t \setminus B_{t-1}) \subseteq X_t$ , and
- (d) for all  $a \in A_{t+1} \setminus A_t$ , the set  $\phi(N(a) \cap B_t)$  is contained in at most  $(\gamma|X_t|)^{\Delta-|N(a) \cap B_t|}$   $V_{t-1}$ -bad  $\Delta$ -tuples in  $X_t$ .

Note that the above properties hold for  $t = 0$  since  $N(a) \cap B_0 = \emptyset$  for all vertices  $a$  (where we define  $B_{-1} = V_{-1} = \emptyset$ ). For some  $t \geq 0$ , suppose that we are given a set  $X_t$  and a map  $\phi$  defined on  $A_t \cup B_t$  that satisfies the above properties. Define  $G_t$  as the subgraph of  $G$  induced on  $V_t = V \setminus \phi(A_t \cup B_t)$ . Since  $|A_t \cup B_t| \leq |V(H)| \leq \delta n$ , the given minimum degree condition on  $G$  implies that  $G_t$  has minimum degree at least  $\alpha n \geq \alpha|V(G_t)|$ . In particular, this implies that  $G_t$  has at least  $\alpha n$  vertices.

Apply Lemma 5.1 to  $G_t$  with  $(X_0)_{5.1} = X_t \setminus \phi(A_t \cup B_t)$ ,  $\beta_{5.1} = 8\beta$ , and  $\alpha_{5.1} = \alpha$  to obtain a set  $X_{t+1}$  satisfying the following properties:

- (i)  $|X_{t+1}| \geq \frac{1}{2}\alpha^{2\Delta}|V(G_t)| \geq \frac{1}{2}\alpha^{2\Delta+1}n$ ,

- (ii)  $|X_t \cap X_{t+1}| \geq \frac{1}{2}\alpha^{2\Delta} |(X_t \setminus \phi(A_t \cup B_t))| \geq \frac{1}{2}\alpha^{2\Delta} (|X_t| - 4\beta n) \geq \frac{1}{8}\alpha^{4\Delta+1}n$ , and
- (iii) the number of  $V_t$ -bad  $\Delta$ -tuples in  $X_{t+1}^\Delta$  is at most  $(\gamma|X_{t+1}|)^\Delta$ ,

Note that Properties (a) and (b) immediately follow.

To extend  $\phi$  to  $A_{t+1} \cup B_{t+1}$ , we first extend  $\phi$  to  $B_{t+1} \setminus B_t$ . We embed vertices in  $B_{t+1} \setminus B_t$  one at a time according to the order given by the labelling. Let  $b \in B_{t+1} \setminus B_t$  be the current vertex where we identify  $b$  with the integer in  $[m]$ . Define  $B[b] = B \cap [b]$  and for each vertex  $a \in A$ , define  $d_b(a) = |N(a) \cap B[b]|$ . We maintain the following three properties while extending  $\phi$ :

- (c')  $\phi(B[b] \setminus B_t) \subseteq X_t \cap X_{t+1}$ ,
- (d1) for all  $a \in A_{t+1} \setminus A_t$ , the set  $\phi(N(a) \cap B[b])$  is contained in at most  $(\gamma|X_t|)^{\Delta-d_b(a)}$   $V_{t-1}$ -bad  $\Delta$ -tuples of  $X_t$ , and
- (d2) for all  $a \in A_{t+2} \setminus A_{t+1}$ , the set  $\phi(N(a) \cap B[b])$  is contained in at most  $(\gamma|X_{t+1}|)^{\Delta-d_b(a)}$   $V_t$ -bad  $\Delta$ -tuples of  $X_{t+1}$ .

Initially, we may assume that  $b = 2t\beta n$  so that  $B[b] = B_t$ . Then Property (c') holds vacuously, and Property (d1) holds by Property (d) of the previous iteration. Moreover, note that if  $a \in A_{t+2} \setminus A_{t+1}$ , then  $a$  is adjacent to a vertex in  $B_{t+2}$ , thus to a vertex with label at least  $2(t+1)\beta n + 1$ . Hence by the definition of bandwidth, it cannot be adjacent to a vertex in  $B_t$ , implying that  $N(a) \cap B_t = \emptyset$ . This implies (d2) at the initial stage, by Property (iii).

Let  $b \in B_{t+1} \setminus B_t$  be the next vertex to embed. Let  $a_1, a_2, \dots, a_d$  be the neighbors of  $b$  (for  $d \leq \Delta$ ). Note that by the definition of  $A_t$ , we have  $a_i \notin A_t$  for all  $i \in [d]$ . On the other hand for each  $i \in [d]$ , since  $b \in B_{t+1} \subseteq [(2t+2)\beta n]$  and  $H$  had bandwidth at most  $\beta n$ , the vertex  $a_i$  cannot be adjacent to a vertex in  $((2t+4)\beta n, m]$ . This implies that  $a_i \in A_{t+2} \setminus A_t$ . For each  $i \in [d]$ , define  $N_i = N(a_i) \cap [b-1]$  and note that  $\phi$  is already defined on  $N_i$ . For each  $i \in [d]$ , since  $a_i$  is adjacent to  $b$  and  $H$  has bandwidth at most  $\beta n$ , the vertex  $a_i$  cannot be adjacent to a vertex in  $[b-2\beta n-1] \cap B \subseteq B_{t-1}$ , thus implying that  $N_i \subseteq B_{t+1} \setminus B_{t-1}$ .

Fix an index  $i \in [d]$ . If  $a_i \in A_{t+1} \setminus A_t$ , then Property (c') implies that  $\phi(N_i) \subseteq X_t$ , and Property (d1) implies that  $\phi(N_i)$  is contained in at most  $(\gamma|X_t|)^{\Delta-|N_i|}$   $V_{t-1}$ -bad  $\Delta$ -tuples of  $X_t$ . Hence there are less than  $\gamma|X_t|$  vertices  $x \in X_t$  for which the  $(|N_i|+1)$ -tuple  $N_i \cup \{x\}$  is contained in more than  $(\gamma|X_t|)^{\Delta-|N_i|-1}$   $V_{t-1}$ -bad  $\Delta$ -tuples of  $X_t$ . If  $a_i \in A_{t+2} \setminus A_{t+1}$ , then Property (c') implies that  $\phi(N_i) \subseteq X_{t+1}$ . Hence similarly as above Property (d2) implies that there are less than  $\gamma|X_{t+1}|$  vertices  $x \in X_{t+1}$  for which the  $(|N_i|+1)$ -tuple  $N_i \cup \{x\}$  is contained in more than  $(\gamma|X_{t+1}|)^{\Delta-|N_i|-1}$   $V_t$ -bad  $\Delta$ -tuples of  $X_{t+1}$ . Since

$$|X_t \cap X_{t+1}| \geq \frac{1}{8}\alpha^{4\Delta+1}n \geq 2\beta n + \frac{1}{16}\alpha^{4\Delta+1}n \geq 2\beta n + d\gamma n,$$

we have  $|(X_t \cap X_{t+1}) \setminus \phi(B[b-1])| \geq |X_t \cap X_{t+1}| - (2\beta n - 1) \geq d\gamma n + 1$  (by Property (c')). Therefore we can choose  $\phi(b) = x$  to maintain Properties (d1) and (d2) by avoiding the vertices identified above for each  $i = 1, 2, \dots, d$ .

Once we finish embedding  $B_{t+1}$ , we greedily embed the vertices  $a \in A_{t+1}$  one at a time. Note that  $\phi(N(a) \cap B)$  is contained in less than  $(\gamma|X_{t+1}|)^{\Delta-|N(a) \cap B|} < |X_{t+1}|^{\Delta-|N(a) \cap B|}$   $V_{t-1}$ -bad  $\Delta$ -tuples. Since the number of  $\Delta$ -tuples containing  $\phi(N(a) \cap B)$  is  $|X_{t+1}|^{\Delta-|N(a) \cap B|}$ , this

in particular implies that there exists a  $V_{t-1}$ -good  $\Delta$ -tuple containing  $\phi(N(a) \cap B)$ . Since every  $V_{t-1}$ -good tuple has at least  $8\beta n$  common neighbors in  $V_{t-1}$ , we thus see that  $\phi(N(a) \cap B)$  has at least  $8\beta n$  common neighbors in  $V_{t-1}$ . By (5), we see that  $|V_t \setminus V_{t-1}| \leq 4\beta n$  and thus  $\phi(N(a) \cap B)$  has at least  $4\beta n$  common neighbors in  $V_t$ . Therefore again by (5), we will never run out of vertices while greedily embedding the vertices in  $A_{t+1}$  to appropriate vertices in  $V_t$ . Note that Property (c) for the next step is satisfied by Property (c'), and Property (d) for the next step is satisfied by Property (d2).  $\square$

Theorem 1.6 has the following interesting corollary which shows that a transference-type result holds even if  $\beta$  is as large as  $c^\Delta$  for some constant  $c$  when the given graph is bipartite and has small bandwidth.

**Corollary 5.2.** *For every positive real number  $\varepsilon$ , there exists a real number  $c < 1$  such that the following holds. If  $G$  is a  $n$ -vertex bipartite graph of maximum degree at most  $\Delta$  and bandwidth at most  $c^\Delta n$ , then  $r(G) \leq (4 + \varepsilon)n$ .*

*Proof.* Define  $c = \frac{1}{256\Delta} \left( \frac{4(4+\varepsilon)}{\varepsilon} \right)^{6\Delta+1}$ . Let  $N = (4 + \varepsilon)n$  and suppose that the edge set of  $K_N$  has been two-colored using red and blue. Without loss of generality, we may assume that the red graph has density at least  $\frac{1}{2}$ . Then we can find a subgraph of the red graph having minimum degree at least  $(1 + \frac{\varepsilon}{4})n$ . Apply Theorem 1.6 to this graph with  $n_{1.6} = N$ ,  $\delta_{1.6} = \frac{1}{4+\varepsilon}$ , and  $\alpha_{1.6} = \frac{\varepsilon}{4(4+\varepsilon)}$  to find a monochromatic copy of  $G$ .  $\square$

## 6 Concluding Remarks

The main theorem of this paper (Theorem 1.4) is a transference principle for Ramsey numbers of bounded degree graphs. It asserts that for all  $\Delta, \xi$  and  $\varepsilon$ , there exists  $\beta$  and  $n_0$  such that the following holds for all  $n \geq n_0$ : if  $G$  is a  $n$ -vertex graph of maximum degree at most  $\Delta$  having a homomorphism  $f$  to  $H$  such that  $|f^{-1}(v)| \leq \beta n$  for all  $v \in V(H)$ , then  $r(G) \leq (1 + \xi)\hat{r}_\varepsilon(H, w) \cdot \beta n$ . Similar result can be proved for more than two colors and for off-diagonal Ramsey numbers using the same approach. The bound on  $\beta$  that we obtain is of tower-type which is unlikely to be best possible. For example, Corollary 5.2 shows that we may take  $\beta \leq c^\Delta$  for some special case.

It might be the case that the transference principle holds for classes of graphs more general than bounded degree graphs.

**Question 6.1.** *Can Theorem 1.4 be extended to degenerate graphs?*

The main difficulty in following the same strategy used in this paper lies in developing a variant of the blow-up lemma that we used. In fact there has been some recent work on extending the blow-up lemma to classes of graphs beyond bounded degree graphs. For an integer  $a$ , a graph is called  $a$ -arrangeable if its vertices can be ordered as  $x_1, \dots, x_n$  such that  $|N(N(x_i) \cap R_i) \cap L_i| \leq a$  for all  $i \in [n]$ , where  $R_i = \{x_{i+1}, \dots, x_n\}$  and  $L_i = \{x_1, \dots, x_i\}$ . Böttcher, Taraz, and Würfl [5] extended the blow-up lemma to arrangeable graphs (after adding a weak constraint on the

maximum degree). Their result implies that a transference-type result holds if the target graph  $H$  is a bounded degree graph. There also has been some partial success towards extending the blow-up lemma to degenerate graphs [22] but only when the bandwidth is small and for almost spanning subgraphs. It is plausible that some of the ideas used in these papers will help answering Question 6.1.

Recall that for a given weighted graph  $(G, w)$ ,  $\hat{r}_\varepsilon(G)$  is not necessarily finite if  $\varepsilon$  is large. In fact  $\hat{r}_\varepsilon(G)$  is finite if and only if  $\varepsilon < \frac{1}{r(\chi(G))-1}$  (where  $r(k)$  is the Ramsey number of  $K_k$ ). Let  $s = r(\chi(G))-1$ . If  $\varepsilon \geq \frac{1}{s}$ , then one can consider a red/blue coloring of  $K_s$  with no monochromatic copy of  $K_{\chi(G)}$  and take a balanced blow-up of this coloring to find an arbitrarily large  $n$ -vertex graph with minimum degree at least  $(1 - \frac{1}{s})n$  having no monochromatic subgraph of chromatic number at least  $\chi(G)$ . In particular, it does not contain a monochromatic copy of  $G$ . On the other hand if  $\varepsilon < \frac{1}{s}$ , then one can show that by supersaturation, for sufficiently large  $n$  there exists  $\Omega(n^{\chi(G)})$  monochromatic copies of  $K_{\chi(G)}$  in every red/blue coloring of an  $n$ -vertex graph  $\Gamma$  of minimum degree at least  $(1 - \varepsilon)n$ . Without loss of generality, assume that at least half of such copies of  $K_{\chi(G)}$  are red. Consider a  $\chi(G)$ -uniform hypergraph over the vertex set of  $\Gamma$  where we place a hyperedge over all red copies of  $K_{\chi(G)}$  in the coloring above. By Kövári-Sós-Turán theorem for hypergraphs, we can find a complete  $\chi(G)$ -partite graph with  $|V(G)|$  vertices in each part if  $n$  is sufficiently large. This implies that we can find a monochromatic copy of  $G$  in  $\Gamma$ .

**Acknowledgements.** I thank David Conlon, Jacob Fox, and Benny Sudakov for fruitful discussions.

## References

- [1] P. Allen, G. Brightwell, and J. Skokan, Ramsey-goodness – and otherwise, *Combinatorica* **33** (2013), 125–160.
- [2] N. Alon and J. Spencer, The probabilistic method, 2nd ed., Wiley, New York (2000).
- [3] J. Böttcher, K. Pruessmann, A. Taraz, and A. Würfl, Bandwidth, expansion, treewidth, separators, and universality for bounded degree graphs, *European J. of Combin.* **31** (2010), 1217–1227.
- [4] J. Böttcher, M. Schacht, and A. Taraz, Proof of the bandwidth conjecture of Bollobás and Komlós, *Mathematische Annalen* **343** (2009), 175–205.
- [5] J. Böttcher, A. Taraz, and A. Würfl, Spanning embeddings of arrangeable graphs with sub-linear bandwidth, arXiv:1305.2078 [math.CO].
- [6] S. Burr and P. Erdős, On the magnitude of generalized Ramsey numbers for graphs, in *Infinite and Finite Sets I*, Colloq. Math. Soc. Janos Bolyai 10, North-Holland, Amsterdam (1975), 214–240.

- [7] V. Chvátal, V. Rödl, E. Szemerédi, and W. Trotter, The Ramsey number of a graph with bounded maximum degree, *J. Combin. Theory B*, **34** (1983), 239–243.
- [8] D. Conlon, Hypergraph packing and sparse bipartite Ramsey numbers, *Combin. Probab. Comput.* **18** (2009), 913–923.
- [9] D. Conlon, J. Fox, and B. Sudakov, On two problems in graph Ramsey theory, *Combinatorica* **32** (2012), 513–535.
- [10] D. Conlon, J. Fox, and B. Sudakov, Recent developments in graph Ramsey theory, arXiv:1501.02474 [math.CO].
- [11] N. Eaton, Ramsey numbers for sparse graphs, *Discrete Math.* **185**, 63–75.
- [12] P. Erdős, Some remarks on the theory of graphs, *Bull. Amer. Math. Soc.* **53** (1947), 292–294.
- [13] P. Erdős and G. Szekeres, A combinatorial problem in geometry, *Compositio Mathematica* **2** (1935), 463–470.
- [14] S. Gerke, Y. Kohayakawa, V. Rödl, and A. Steger, Small subsets inherit sparse  $\varepsilon$ -regularity, *J. Combin. Theory B*, **97** (2007), 34–56.
- [15] R. Graham, V. Rödl, and A. Ruciński, On graphs with linear Ramsey numbers, *J. Graph Theory* **35** (2000), 176–192.
- [16] R. Graham, V. Rödl, and A. Ruciński, On bipartite graphs with linear Ramsey numbers, *Combinatorica* **21** (2001), 199–209.
- [17] R. Graham, B. Rothschild, and J. Spencer, Ramsey theory, Wiley, New York (1990).
- [18] J. Fox and B. Sudakov, Density theorems for bipartite graphs and related Ramsey-type results, *Combinatorica* **29** (2009), 153–196.
- [19] J. Fox and B. Sudakov, Two remarks on the Burr-Erdős conjecture, *European J. Combin.* **30** (2009), 1630–1645.
- [20] J. Komlós, G. Sárközy, and E. Szemerédi, Blow-up Lemma, *Combinatorica* **17** (1997), 109–123.
- [21] J. Komlós and M. Simonovits, Szemerédi’s regularity lemma and its applications in graph theory, Bolyai Society Mathematical Studies 2, Combinatorics, Paul Erdős is Eighty (Volume 2) (D. Miklós, V. T. Sós, T. Szőnyi eds.), Budapest (1996), 295–352.
- [22] C. Lee, Embedding degenerate graphs of small bandwidth, arXiv:1501.05350 [math.CO].
- [23] L. Lovász, On decomposition of graphs, *Studia Sci. Math. Hungar.* **1** (1966), 237–238.
- [24] F. Ramsey, On a problem of formal logic, *Proc. London Math. Soc.* **30** (1930), 264–286.